Understanding Developers Well-being and Productivity: A 2-year Longitudinal Analysis during the COVID-19 Pandemic—RCR Report

DANIEL RUSSO, Department of Computer Science, Aalborg University, Denmark
PAUL H. P. HANEL, Department of Psychology, University of Essex, United Kingdom
NIELS VAN BERKEL, Department of Computer Science, Aalborg University, Denmark

The artifact accompanying the paper "Understanding Developers Well-Being and Productivity: A 2-year Longitudinal Analysis during the COVID-19 Pandemic" provides a comprehensive set of tools, data, and scripts that were utilized in the longitudinal study. Spanning 24 months, from April 2020 to April 2022, the study delves into the shifts in well-being, productivity, social contacts, needs, and several other variables of software engineers during the COVID-19 pandemic. The artifact facilitates the reproduction of the study's findings, offering a deeper insight into the systematic changes observed in various variables, such as well-being, quality of social contacts, and emotional loneliness. By providing access to the evidence-generating mechanisms and the generated data, the artifact ensures transparency and reproducibility and allows researchers to use our rich dataset to test their own research question. This Replicated Computational Results report aims to detail the contents of the artifact, its relevance to the main paper, and guidelines for its effective utilization.

CCS Concepts: • **Social and professional topics** → *Computing industry*; **Management of computing and information systems**; **Project and people management**;

Additional Key Words and Phrases: COVID-19, longitudinal study, well-being, future of work, IJARS Model

ACM Reference Format:
Daniel Russo, Paul H. P. Hanel, and Niels van Berkel. 2024. Understanding Developers Well-being and Productivity: A 2-year Longitudinal Analysis during the COVID-19 Pandemic—RCR Report. *ACM Trans. Softw. Eng. Methodol.* 33, 3, Article 82 (March 2024), 4 pages. https://doi.org/10.1145/3640338

1 OVERVIEW

The paper presents a longitudinal study conducted over 24 months, from April 2020 to April 2022, exploring the changes in well-being, productivity, social contacts, needs, and several other variables of software engineers during the COVID-19 pandemic. The research provides insights into the systematic shifts observed in various aspects of developers' lives and work during this period.

Table 1. Overview of artifacts, location, and description

Artifact	Location (DOI)	Description
Data Files	https://zenodo.org/records/10778925	Contains data for waves 1 through 6 and correlation data.
Survey Files	https://zenodo.org/records/10778925	PDFs representing the questionnaires for each wave.
Analysis Script	https://zenodo.org/records/10778925	R script detailing the analysis performed.

The paper posits that the COVID-19 pandemic has led to significant shifts in work arrangements, impacting the well-being and productivity of software engineers. We found that as lockdown measures were relaxed, there was an observable increase in well-being and the quality of social contacts, while emotional loneliness decreased. Despite these changes, boredom and productivity levels remained stable throughout the study period. The research also introduces the Integrated Job Demands-Resources and Self-Determination Model as a framework to explain these observations.

The findings of the study indicate that developers showed a preference for hybrid work arrangements post-pandemic. This highlights the evolving nature of work in the software engineering domain. Additionally, the study found that prior job changes and low job satisfaction were consistently linked to intentions to change jobs if the current work conditions did not meet developers' needs. This underscores the importance for software organizations to adapt to various work arrangements to remain competitive employers in a post-pandemic era. For a detailed discussion, we refer the reader to the Reference [1].

1.1 Artifacts

The artifacts are available on Zenodo and consist of three main components—as highlighted in Table 1.

The artifact includes the raw data files from all six waves of the study. Additionally, survey files in PDF format represent the questionnaires administered during each wave to collect primary data from the participants.

The data files contain empirical results derived from the study, including data from each wave and correlation data. This evidence provides insights into the well-being, productivity, and preferences of software engineers during the pandemic.

The provided R script, 'Six wave longitudinal study.r', details the analysis performed on the dataset. To replicate the study's findings, one can run this script, ensuring that all data files and the script are in the same working directory.

1.2 Prerequisites and Requirements

No specific hardware or OS requirements are needed for the replication of the study. However, for the analysis, a computer with R installed is necessary to run the provided R script. Additionally, a PDF reader is required to view the survey files.

The targeted population for this study was software engineers. The participants were selected from a pool of 483 software engineers, and the study eventually included 192 professionals who met specific criteria related to their work environment during the COVID-19 pandemic.

Participants were pre-screened based on their knowledge of software development techniques, their profession as computer programmers, their use of technology at work, and their reliability score on the Prolific platform. A competence screening was conducted to ensure participants had

adequate knowledge of software development. Quality screening was also implemented to ensure the reliability of the data. For this specific study, participants should not have lived in countries with non-uniform COVID regulations and should have been working from home during the pandemic for at least 50% of their time.

To replicate the study, one should

(1) familiarize themselves with the survey content by reviewing the provided survey PDFs,
(2) examine the data in the provided CSV files,
(3) ensure that R is installed on their computer,
(4) place all data files and the R script in the same working directory, and
(5) run the R script 'Six wave longitudinal study.r' to replicate the analysis.

1.3 Steps to Reproduce

The data for this study were collected through a longitudinal survey spanning 24 months with six measurement points. The surveys were distributed to software engineers, and the responses were collected over the specified period. To generate similar data, one would need to

(1) identify a pool of software engineers fitting the criteria mentioned in the "Participants" section.
(2) distribute the surveys (provided in the replication package as 'Survey wave 1.pdf' to 'Survey wave 6.pdf') at regular intervals, and
(3) collect and anonymize the responses, ensuring identifiers are removed for privacy.

The data analysis for this study was performed using an R script provided in the replication package. To reproduce the analysis, follow these steps:

(1) Ensure you have R installed on your computer.
(2) Download the data files ('Waves 1-6 without ID.csv', 'Country1-6.csv', and correlation files) and the R script ('Six wave longitudinal study.r') from the replication package.
(3) Place all the downloaded files in the same working directory.
(4) Open the R script 'Six wave longitudinal study.r' in R or any R-compatible IDE.
(5) Run the script to perform the analysis. The script will process the data, perform the necessary statistical tests, and generate the results.

2 ARTIFACT INSTRUCTIONS

2.1 Preparing the Artifact

To prepare the artifact for replication and analysis, follow the following steps:

(1) *Download the replication package:*
    — Navigate to the provided Zenodo link: https://zenodo.org/records/10778925.
    — Download the entire replication package, which will include data files, survey files, and the analysis script.
(2) *Organize the files:*
    — Create a dedicated folder on your computer named "Replication_Package_COVID19_Study".
    — Extract the downloaded files into this folder, ensuring that all files are organized in a manner that's easy to navigate.
(3) *Software requirements:*
    — Ensure you have R installed on your computer. If not, then download and install R from the official website: https://cran.r-project.org/.

— Optionally, you can also install RStudio, an integrated development environment for R, from https://www.rstudio.com/.

(4) *Familiarize yourself with the data:*
    — Open and review the data files, especially 'Waves 1-6 without ID.csv' and 'Country1-6.csv', to understand the structure and content of the dataset.
    — Review the correlation files to understand the relationships between different variables across the six waves.

(5) *Review the surveys:*
    — Go through the survey PDFs ('Survey wave 1.pdf' to 'Survey wave 6.pdf') to understand the questions posed to the participants and the nature of the data collected.

(6) *Set up the analysis environment:*
    — Ensure that the R script ('Six wave longitudinal study.r') and all data files are in the same working directory.
    — Open the R script in R or RStudio to review its content and understand the analysis steps.

By following these steps, you will have successfully prepared the artifact for replication and analysis. The next section will guide you through the actual replication process.

REFERENCES

[1] Daniel Russo, Paul H. P. Hanel, and Niels van Berkel. 2023. Understanding developers well-being and productivity: A 2-year longitudinal analysis during the COVID-19 pandemic. *ACM Trans. Softw. Eng. Method.* (2023).